

DOCUMENT RESUME

ED 202 202

FL 012 137

AUTHOR Marckworth, M. Lois
TITLE An Experimental Investigation of Verbs of Judging in the Subjective Lexicon in Japanese and English.
INSTITUTION Linguistic Research, Inc., Edmonton (Alberta).
REPORT NO ISBN-0-88783-022-6
PUB DATE 78
NOTE 11p.; Paper presented at the Annual Meeting of the (Canadian) Western Conference on Linguistics (8th, October 20-21, 1978). In its Proceedings of the Eighth Annual Meeting of the Western Conference on Linguistics, p90-98, 1978.
AVAILABLE FROM Linguistic Research Inc., P.O. Box 5677, Station L, Edmonton, Alberta, Canada T6C 4G1 (\$10.00 for entire proceedings).
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cross Cultural Studies; *English; English (Second Language); *Japanese; *Semantics; Statistical Analysis; *Verbs
IDENTIFIERS Hierarchical Cluster Structure; *Lexical Semantics

ABSTRACT

Lexical semantics deals with the determination and description of the semantic knowledge or structure embodied in the lexicon. A description of lexical structure focuses on the meaning of a single lexical item in terms of its place in a structure consisting of related lexical items. This study concerns Japanese and English verbs of judging with regard to the formulation of a methodology for the process of making determinations. These verbs are interpersonal relationship verbs in that they reflect a judge's feelings in regard to a defendant's acts. The statistical analysis technique was employed to capture clues to the semantic dimensions of four groups of Japanese and English verbs: one group per language of negative verbs of judging and one group per language of positive verbs of judging. The results are presented in charts depicting the hierarchical cluster structure of the verbs in each set. The meaning similarities between the members of each set are reflected in cohesion scores that define the intensity of the bonding of a term into the hierarchical structure. (JK)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

Proceedings of the Eighth Annual Meeting of the
Western Conference on Linguistics (October 20-21, 1978)

AN EXPERIMENTAL INVESTIGATION OF
VERBS OF JUDGING IN THE SUBJECTIVE
LEXICON IN JAPANESE AND ENGLISH

M. Lois Marchworth

Department of Linguistics, University of Alberta, Edmonton, Canada

The study that I would like to talk about today is in the area of lexical semantics, and is an initial report on a project that is just beginning to develop. There is a long tradition of interest among linguists in (1) how to determine and (2) how to describe the semantic knowledge (or structure) presumed to be embodied in the lexicon; that is, in how to find out about and characterize the shared knowledge that allows a group of people to use lexical items with some assurance of transmitting accurate meaning between speaker and hearer.

A number of different techniques have been used in the description of lexical structure—such techniques as combinatorial analysis, semantic feature attribution, semantic network construction, computer-implemented artificial intelligence modeling, etc., and all of these techniques attempt to characterize, in some terms or other, the relations between lexical items. It seems axiomatic in any endeavour of this sort that the meaning of a lexical item is most interestingly viewed through its place in a structure of other related lexical items—a point of view which is acknowledged in the study reported here.

However, despite the range of techniques available for the description of lexical structure, it is my impression that the logically prior activity, the determination of such structure, has largely been dependent on introspection or on the elicitation of the intuitions of a native-speaking informant. The study reported here is chiefly concerned with a more rigorous methodology for making such determinations, with the description of them emerging largely as a by-product of the statistical analytical technique.

We might begin with a sketch of the technique and some comments on its value. The technique is an experimental one and was developed in its present form by Fillenbaum and Rapoport (1971), who discuss it as a tool for investigating the structures of the subjective lexicon although a better term might be the intersubjective lexicon, since it purports to elicit shared knowledge from a pool of speakers concerning word meanings.

The technique uses as stimulus materials a set of semantically-related lexical items, and we will return shortly to the problem posed by the selection of such a set since it is critical to the results obtained and by no means as easy as might be supposed. An example of such a set might be the following temperature terms: freezing, cold, tepid, warm, hot, boiling, etc. Subjects are asked, on an individual basis, to make judgements about the similarity and difference of the words in the set compared by pairs. (The actual presentation is by a computer-generated upper triangular matrix in which each item in the set is compared with every other item in the set; the presentation order is randomized individually for each subject to control for order effects and fatigue.) Subjects express their similarity/difference judgements in terms of a nine-point scale on which 1 indicates "most similar" and 9 indicates "most different." As a first step, subjects are asked to scan the whole set and choose the pair of terms they judge to be most similar in meaning and the pair they judge to be the most different, assigning a 1 and a 9 respectively to these pairs. In terms of our exemplar set, a 9 might be assigned to the pair freezing:boiling and a 1 to freezing:cold (or perhaps to hot:boiling or warm:tepid). This initial step sets anchor points for the ends of the scale and encourages subjects to make use of the whole scale in assigning ratings; there is no limit to how many 1's and 9's a subject may assign, only the constraint that he must have at least one of each. Subjects then complete the judgements on the whole set, and these numbers are

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

A. L. Vanek

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

2 * This document has been reproduced as
received from the person or organization
originating it.

Minor changes have been made to improve
reproduction quality.

* Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

ED202202

FL012137

the raw data for the statistical analysis.

Several sorts of statistical analysis may be performed on these elicited judgements with a view to identifying significant inter-subject trends, which, in turn, will provide clues to the bases upon which subjects made their judgements. Such bases are assumed to be the semantic dimensions or constructs of the lexical set under consideration, upon which speakers and hearers manipulate lexical choices during the production or comprehension of an utterance. Hierarchical clustering is the technique used in the study reported here, although if a lexical set exhibits clear dimensionality, multidimensional scaling may be a more informative technique. For example, our exemplar set yields a single strong dimension—that of degree of heat—under a multidimensional scaling analysis. More complex lexical sets usually turn out to have several such dimensions (for example, colour terms analysed in this fashion yield three dimensions: hue, saturation, and lightness), or to be more usefully described in terms of hierarchies of sub-clusters of items.

The value of such a technique for the investigation of lexical semantics seems to me to be based on four considerations. (1) It is an experimental technique, and the results are therefore subject to validation by replication across studies. (2) It elicits a simple and direct judgement about meaning relationships from the subject, rather than asking him to provide an introspective verbalization about meanings and meaning differences in a lexical set—a task which is difficult for even the most verbally sophisticated speaker. (3) As Fillenbaum and Rapoport point out, the technique allows an a posteriori analysis of subjects' responses rather than suggesting an a priori analysis and asking subjects for agreement or disagreement. (Lawyers call this latter technique "leading the witness;" it should be avoided in lexical study as well as in court.) (4) The technique serves as a seedbed for further investigations, both because it identifies subtle and promising trends in the data and because studies can be repeated, varying the lexical set or the subject pool along principled lines.

It seemed possible that the use of this technique might provide entrée to the area of cross-cultural semantic studies, allowing access to the question of whether comparable lexical domains in different languages have comparable semantic dimensions. Such a question, of course, is of both theoretical and practical interest. In theoretical terms, one may wonder whether such semantic dimensions are linguistic universals or whether they are largely culturally determined. In practical terms, the description of lexical relations is a continuing problem in the teaching of a second language; the poorer the understanding of a topic, the poorer the teaching of it, and lexical semantics has long been a poor relation of phonology, morphology, and syntax.

The specific study in which this technique was utilized, and which I wish to report on today, concerned a subset of the verbs of judging from the lexicons of both Japanese and English. Verbs of judging may be used with an agent and a patient, usually referred to as the Judge and the Defendant respectively. They express the Judge's feelings about an act of the Defendant, and thus are part of the larger category of interpersonal relationship verbs. The verbs of judging were chosen for the study for two reasons: (1) work on them in English had already been done, and we hoped both to replicate and to extend these studies (Fillenbaum and Rapoport 1971, Magners 1977); (2) we felt that lexical items from the domain of interpersonal relationships might be more sensitive to cultural differences than a set, such as heat or colour terms, that more directly referred to the physical universe.

The lexical sets that were tested are shown in Table I. You will note that in both English and Japanese, the verbs are divided into positive and negative terms, thus making four test sets in all. Fillenbaum and Rapoport (1971) tested a similar, although somewhat smaller, set of English terms with the positive and negative items pooled. (Subjects can only deal with about 20 terms in this task; we were thus able to increase the

number of terms investigated by dividing them into two sets.) Magnera (1977), as part of a larger study, replicated Fillenbaum and Rapoport's findings, changing the set membership slightly but revealing the same semantic structure. It was she who suggested that the set should be divided into positive and negative subsets; both she and Fillenbaum and Rapoport had found the strongest semantic dimension in the structure to be a good:bad judgement scale, and she felt that this very strong dimension might be masking other pertinent information.

The Japanese sets were selected by M. Kawashima, a graduate student in our department. In addition, she carried out the testing in Japan and assisted in the interpretation of the results. For the convenience of English speakers, she has provided approximate translations; however, the reader should keep in mind that his use of these translations will sum up the semantic dimensions of English rather than those of Japanese.

ENGLISH: Positives

1. accept
2. acclaim
3. acquit
4. approve of
5. commend
6. condone
7. excuse
8. forgive
9. justify
10. honor
11. pardon
12. praise

Negatives

1. accuse
2. blame
3. censure
4. chide
5. convict
6. criticise
7. denounce
8. deride
9. disapprove of
10. discredit
11. reject
12. reproach
13. scold

JAPANESE: Positives

1. bengosuru 'justify'
2. homeru 'commend'
3. jimansuru 'brag (be proud) of'
4. kabansuru 'forgive (a)'
5. minaosu 'commend in surprise'
6. minogasu 'overlook indulgently'
7. mitomeru 'approve openly'
8. mokuninsuru 'approve tacitly'
9. shijisuru 'support'
10. sonkeisuru 'respect'
11. suuhaisuru 'adulate'
12. tataeru 'praise'
13. ukeireru 'accept'
14. yurusu 'forgive (b)'
15. zessansuru 'laud'

Negatives

1. azawaru 'laugh at scornfully'
 2. bakanisuru 'ridicule'
 3. hihansuru 'criticise'
 4. hinansuru 'censure'
 5. keebetsusuru 'be contemptuous of'
 6. kenasu 'speak ill of (a)'
 7. kobamu 'reject'
 8. kusu 'speak ill of (b)'
 9. mikagiru 'wash one's hands of someone'
 10. mikudasu 'look down on'
 11. najiru 'reproach (a)'
 12. semeru 'reproach (b)'
 13. shikāru 'scold'
 14. shiseekisuru 'reprimand'
 15. togameru 'blame'
-

TABLE I. The four test sets used in the study.

Choice of the lexical items to be included in the test set is perhaps the most difficult part of the experimental technique, and it is important during the interpretation of the results to remember that conclusions, in the last analysis, are valid for the set tested and the relationships that obtain within it. As noted, subjects in the experimental procedure can cope with about twenty terms in a set, so that if the semantic field being investigated is clearly bounded and contains a small number of terms, the set may be comprehensive. However, if the semantic field is very large or has fuzzy boundaries, choices must be made—and preferably on some principled grounds. The domain of the verbs of judging, although fairly clearly defined, is very large in both English and Japanese. Choice of items was made on the grounds of (1) the attempt to choose a representative sample, avoiding what the experimenters felt to be semi-synonyms, (2) the avoidance of any very low frequency, archaic, or ambiguous terms, and (3) for the English sets, only, the use of the basic set—with a few alterations—used by Fillenbaum and Rapoport and by Magnera. This latter constraint was imposed because the present study hoped (1) to show the replicability of the earlier studies and (2) to extend Magnera's study along the lines she suggested, suppressing the strong, masking good:bad judgement dimension. Set members were selected with the full realization that the addition of a new term or the loss of an old would possibly change the equilibrium of the inter-related set and yield a new semantic structure; this problem was, of course, more acute with the Japanese sets than with the English since the latter were partial replications and the former a first attempt. The results should be evaluated accordingly.

Subjects doing the experiment were asked to place the verbs in a sentence frame when comparing them: They (verb) him for English and for Japanese, a frame that is best translated 'People (verb) that person.' This constraint was imposed to provide subjects with a neutral and standard semantic context and to direct their attention to the meaning of the verb rather than allowing them to focus on graphemic shape, sound, or some other idiosyncratic variable. Items in the sets were randomized individually in each (computer produced) presentation matrix to control for order effects and subject fatigue. In the case of the English sets, different groups of subjects were used for the positive and negative sets; 24 were used for the positive set and 20 for the negative. For the Japanese sets, every subject did both a positive and a negative matrix, of which 144 positive and 143 negative matrices were used in the statistical analysis.

Analysis of the data was attempted using both the hierarchical clustering technique and multidimensional scaling. Of these, the hierarchical clustering provided the most interpretable results, with the multidimensional scaling suggesting some interesting trends which may repay further investigation.

The structures produced by the hierarchical clustering analysis for each test set may be seen in Figures 1-4. These structures should be read as follows. The figures on the ordinate are cohesion scores which signal, by their relative value, the intensity of the bonding of a term into the hierarchical structure; e.g., in Figure 1, displaying the English positive structure, acclaim and praise have a cohesion score of about .16 and accept and approve of a cohesion score of about .31, indicating that subjects judged acclaim and praise to be more similar in meaning than they did accept and approve of. Thus the lower down a verb coheres into a cluster, the more tightly bound to it it is in terms of judged meaning similarity. You will notice, again in Figure 1, that the English positive structure contains three major subclusters: the first is made up of justify, acquit, excuse, forgive, and pardon; the second of accept, approve of, and condone; the third of commend, honor, acclaim, and praise. Consider the first subcluster: forgive and pardon form the nucleus as the most cohesive (judged most similar) pair; excuse is next added to the cluster by assessing it against the combined score of the first pair; acquit is added to the first three, and last justify to the whole structure.

The following interpretations may be made of the English positive cluster structure.

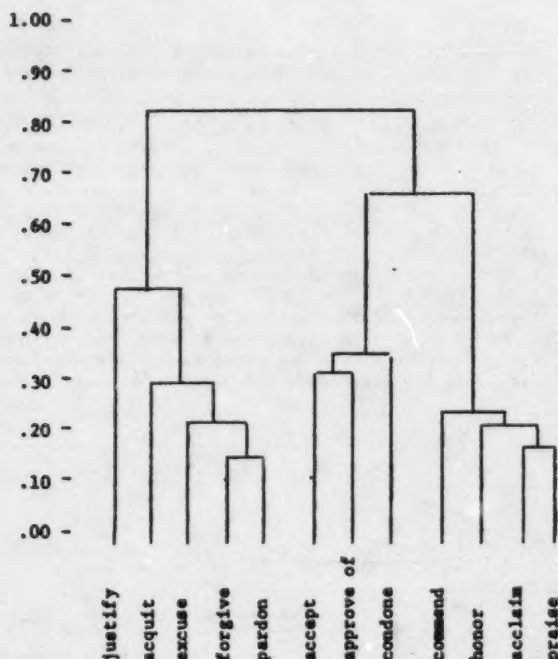


Figure 1. Hierarchical cluster structure for the English positive verbs of judging.

The verbs in the first major subcluster (justify, acquit, excuse, forgive, and pardon) have in common the presupposition of a bad act on the part of the Defendant which is accepted by the Judge, and to some extent excused, overlooked, or actively defended by him. This first cluster is only weakly associated with the other two major subclusters, both of which have members which imply a good act on the part of the Defendant with a positive response on the part of the Judge. The difference between these latter sub-clusters is intensity of this positive response. The weaker of the two (accept, approve of, and condone) is neutral or mildly positive in response, with the implication that the act, itself, was neutral or even mildly negative (e.g., condone). The stronger of the two (commend, honor, acclaim, and praise) suggests very good behaviour with a strongly positive response by the Judge.

Both Fillenbaum and Rapoport and Magners identified two clusters in a similar set of English positives; these two clusters corresponded to cluster one and clusters two and three together from Figure 1 above. The inclusion, in both of the earlier studies, of negative verbs in the test set, with the resultant emergence of an over-riding positive:negative dimension, obscured the less robust but still significant distinction between clusters two and three in this study—the neutral:enthusiastic dimension. The

appearance of this latter dimension justifies Magnera's suggestion of separating the positive and negative verbs into different test sets to suppress the dominant, and thus masking dimension.

The hierarchical cluster structure of the English negative verb set is shown in Figure 2. Note that in this case, also, three major subclusters emerge. The first, which consists of accuse, blame, and convict, suggests a fairly serious bad action on the part of the Defendant and a fairly permanent, or even legalistic, judgement made of it. The second subcluster (deride, discredit, denounce, and reject) carries a strong implication of social rejection by the Judge of the Defendant. The infrastructure of this subcluster shows deride and discredit paired together, and denounce and reject; the first of these pairs has an implication of scorn which is lacking in the second. Subcluster three (censure, reproach, criticise, disapprove of, chide, and scold) also shows an infrastructure of pairs linked together into a larger group: chide and scold are judged very similar in meaning, criticise and disapprove of only a bit less so, and censure and reproach form the weakest pair. We might assume that these pairs fall along a cline of formality or social distance between Judge and Defendant; the group as a whole seems to imply a single

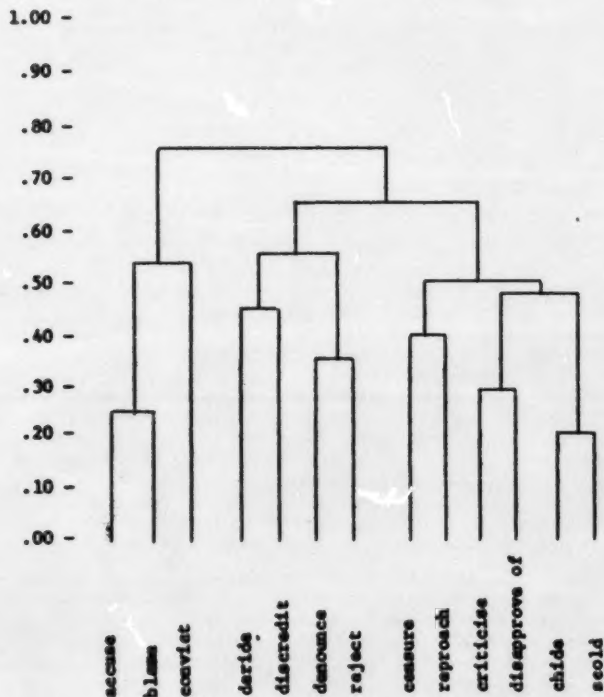


Figure 2. Hierarchical cluster structure for the English negative verbs of judging.

bad action or temporary attitude for which a single reproach is being uttered.

Both Fillenbaum and Rapoport and Magnara, in their earlier studies, found two clusters in the negative verbs (a "disapprove" cluster and a "reject" cluster). This indicates again that the strong negative:positive dimension in the combined set masked subtler structures in the data.

Before we go on to discuss the Japanese results, it is worth noting that, as a whole, the English negative set exhibits a more diffuse structure than the positive. This is demonstrated by the fact that the clusters in the negative set cohere at a higher level.

The hierarchical cluster structure of the Japanese positive test set may be seen in Figure 3. You will note that there are four subclusters that emerge, although we may

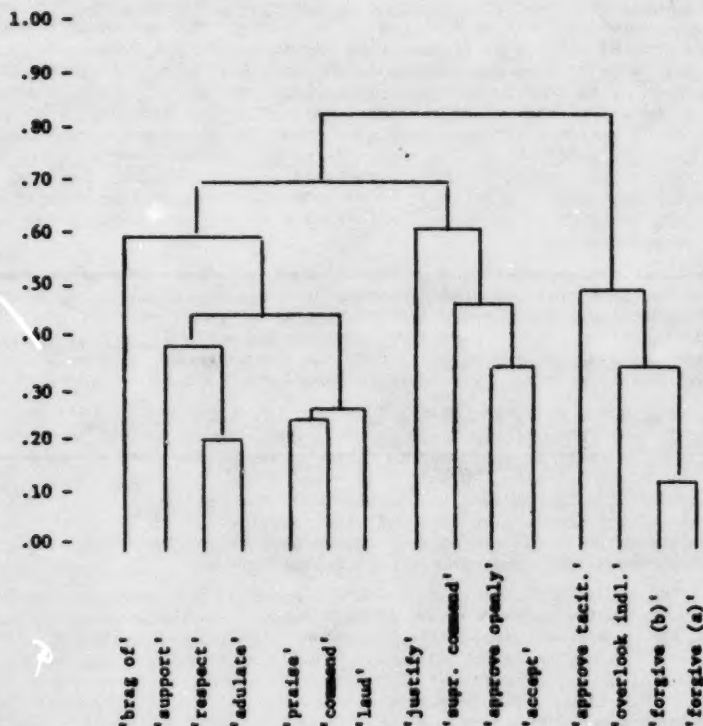


Figure 3. Hierarchical cluster structure for the Japanese positive verbs of judging.

wish to group the first two subclusters together. In addition, note that one item, 'brag of' is an outlier, only weakly connected to anything else. Mrs. Kawashima, who carried out the testing in Japan, has supplied the following comments on the cluster structure of the positives, and the negatives in the following section.

In both the first subcluster ('support', 'respect', and 'adulate') and the second ('praise', 'commend', and 'laud'), the Judge is strongly impressed by a good action on the part of the Defendant. The verbs in the second subcluster always indicate an overtly expressed commendation, whereas those in the first may indicate feelings as well. Within the first subcluster, 'respect' and 'adulate' are used of a Defendant of higher status than the Judge—the first of a teacher or parent and the second, which may also be translated 'worship' of a hero-figure such as an astronaut or a film star. The third member of the subcluster, 'support' is used when the Defendant and the Judge are members of the same crowd. In the second subcluster, the three verbs are very similar in indicating direct praise, but vary in intensity.

The third subcluster ('justify', 'commend in surprise', 'approve openly', and 'accept') indicate much weaker approval on the part of the Judge. The Defendant's action may not have been very commendable, or if it was, this has surprised the Judge. The pair 'approve openly' and 'accept' show the strongest cohesion, both have the implication that the Defendant's behaviour is just so-so—ordinary rather than good. 'Justify' shows the weakest linkage to the group, and it has the implication that the Defendant's behaviour has been poor but the Judge must say something good about it. Subcluster four, which contains 'approve tacitly', 'overlook indulgently', 'forgive (a)' and 'forgive (b)', implies behaviour on the part of the Defendant that is not good at all—behaviour to which the Judge is just closing his eyes. 'Approve tacitly' probably indicates the best of the behaviours in this set, but even in this case the approval is very weak and is not at all a signal of public acceptance.

The hierarchical cluster structure of the Japanese negative verbs of judging is shown in Figure 4. In this case, we find a pattern of two large subclusters, each of which may be divided into two further subclusters, and in addition two outliers which are only weakly connected to the set. The outliers are, on the one end, 'reject' and, on the other end, 'to wash one's hands of someone'; these, in the judgement of the subjects, lacked strong semantic ties to at least this subset of the Japanese negative verbs of judging.

The first large subcluster, containing 'censure', 'reproach (b)', 'reprimand', 'criticise', 'scold', and 'blame', implies a rather formal disapproval of the Defendant's action; emotion, insofar as it is expressed in these verbs, is for the Defendant's behaviour rather than for him personally. Within this subcluster, 'censure', 'reproach (b)' and 'reprimand' cohere closely together on the one hand, and 'criticise', 'scold', and 'blame' do the same on the other. The first of these carries the implication of a judgement within the framework of a failure in duty on the part of the Defendant; the second is composed of more common, and thus perhaps more general verbs.

The second large subcluster contains verbs which convey more emotion than do those in the first subcluster. The infrastructure of this second subcluster contains two quite distinct groups. The first, composed of 'look down on', 'laugh at scornfully', 'ridicule' and 'be contemptuous of', may be called the social ridicule group; they mirror a common social technique for conveying disapproval of a Defendant's action and attempting to force a change. Verbs in this group are only used when the Judge is of higher status than the Defendant. The second group ('reproach (a)', and 'speak ill of (a) and (b)') are emotionally loaded verbs. They imply that the Judge has an emotional involvement with the issue (and perhaps also with the Defendant) and that this is conveyed to the Defendant. These are verbs that are quite likely to provoke an indignant response from the Defendant.

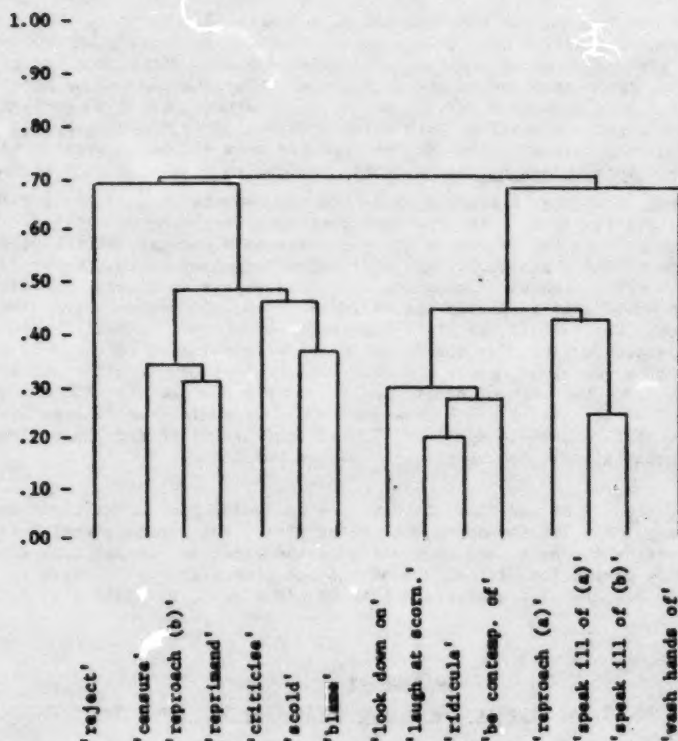


Figure 4. Hierarchical cluster structure for the Japanese negative verbs of judging.

Several points may be made about the results of this study. First, the semantic dimensions for English obtained by Fillenbaum and Rapoport and by Magnera are also obtained in this study, despite changes and additions in the test sets. Such a replication both establishes these dimensions as more robust experimental results, and enhances our respect for the experimental technique. Second, the additional English dimensions predicted by Magnera to have been masked by the dominant positive:negative dimension of the pooled set in her study and in Fillenbaum and Rapoport's, did, indeed, materialize. This finding not only justifies Magnera's specific prediction, but also, and more generally, shows that principled manipulation of the test set provides a technique for more extensive (or intensive) investigation of a lexical domain along lines suggested by a first trial.

Third, the utility of the experimental technique for cross-language comparison appears to be worth pursuing in the light of the interesting similarities and differences

that emerge between the English and the Japanese sets. Note, for example, that both the English and the Japanese positive sets show a major dimension evaluating the Defendant's action as either praiseworthy or as bad-but-forgivable; however, within the two praiseworthy subcategories, Japanese shows an infrastructure reflecting the status levels of Judge and Defendant, and English a neutral:enthusiastic dimension. A comparison of the English and Japanese negatives shows in both cases a verb cluster implying social rejection or ridicule, but the main dimension in the Japanese test set is apparently between emotion-laden and emotionally neutral verbs, while in English this construct is lacking.

Last, it seems to me that a study such as the one reported here just reveals the tip of the iceberg, and leaves us with more to investigate than when we started. For example, we would like to know the degree of group cohesiveness amongst the subjects in the experiment, in terms of their similarity and difference judgements; i.e., we would like to know how potent are the semantic dimensions uncovered experimentally in facilitating lexical choice. We would like to pursue the validity of the dimensions using other experimental techniques; for example, we might construct semantic-differential-type scales using the dimensions and require that the verbs in the original test set as well as others be placed along them, or we might choose a prototypic verb for each cluster and ask for distance judgements from that verb of other cluster members and new test items. When the reliability of a set of dimensions has been established, we would then like to use that set to test subject differences—in age, sex, educational level, social class, language sensitivity, bilingual balance, and many other speaker variables.

Acknowledgements. I should like to thank the following people for their contributions to this study: Michiko Kawashima, for constructing the Japanese test sets, carrying out the testing in Japan, and commenting on the Japanese hierarchical cluster structures; Dr. Wm J. Baker, for statistical advice and consultation; Dr. Gary D. Prideaux and Georgia Magnera, for useful comments on the first presentation of this paper.

References

- FILLENBAUM, S. & RAPOPORT, A. Structures in the Subjective Lexicon. New York: Academic Press. 1971.
- MAGNERA, Georgia. Organization of the Subjective Lexicon. University of Alberta, Department of Linguistics, M.Sc. Thesis. Edmonton, Alberta. 1977.